Guanwen Qiu

Email : guanwenq@seas.upenn.edu

Education

- **University of Pennsylvania** — Philadelphia, PA
  Master of Science in Engineering in Computer and Information Science, GPA: 3.7 — *May 2023 - Dec 2024 (Expected)*
- **University of Pennsylvania** — Philadelphia, PA
  Master of Science in Engineering in Data Science, GPA: 4.0 — *Jan 2022 - May 2023*
- **University of California, Irvine** — Irvine, CA
  Bachelor of Computer Science (Cum Laude), Major GPA: 3.9 — *Sep 2016 - Dec 2020*

Publication

- **Complexity Matters: Dynamics of Feature Learning in the Presence of Spurious Correlations**
  GuanWen Qiu, Da Kuang, Surbhi Goel
  Will be presented as a poster at ICML 2024

Research Experience

- **Understanding representation learning through Boolean function** — UPenn, PA
  Independent Research with Prof. Surbhi Goel — *May 2023 - Current*
  - Propose a novel theoritical setting with accompanied dataset that encapsulate feature as Boolean functions with different complexity and varying correlation to study representation learning of neural network in presence of spurious correlation
  - Conducting carefully defined experiments on neural network under the framework to elucidate critical aspects of deep learning, such as simplicity bias, early learning, and the grokking phenomenon.
  - Study the influence of spurious feature on the representation learned by neural network under SGD both empirically and theoritically.
  - Compare existing debiasing algorithm, such as Rich feature construction, robust loss and re-weighting methods. Point out the weakness of popular spurious dataset and debiasing algorithms.
  - Apply tools in Boolean function analysis to do theoritical analysis and design an algorithm to solve the spurious learning task under the proposed setup.

- **Music Generation through Deep Learning** — UPenn, PA
  Independent Research — *Jan 2022 - June 2022*
  - Compared the performance of different deep learning architectures in generating music MIDIs including Transformer, LSTM, CNN and Variational Auto Encoder (VAE) using Torch and LaKh MIDI Dataset.
  - Designed a highly flexible music midi generation architecture combining LSTM and VAE CNN to produce 5 tracks (Drums, Strings, Piano, Guitar and Bass) MIDI with long term musical development. The model also can be used by musicians to do style transfer and melody generation.
  - Used the pre-trained novel generative model as a downstream model to classify MIDIs into 10 different genres with higher than 80% accuracy.

- **Cell Universe** — UCI
  Research Assistant with Prof. Wayne Hayes — *June 2020 - Jan 2021*
  - Improved the accuracy of bacteria tracking algorithm by producing more realistic synthetic bacteria images.
  - Developed an iterative rendering algorithm to simulate the gray scale light diffraction pattern of different types of bacteria images. Applied gradient descent to tune background color, cell color and diffraction parameters of synthetic bacteria images.
  - Designed a heuristic that consider factors including area of overlap cells, cells orientation and movements to improve the performance and efficiency of the simulated annealing algorithm in aligning synthetic cells with cells in real images.
  - Designed an adaptive temperature scheduling strategy for the simulated annealing algorithm.

Teaching experience

- TA CIS522: Deep Learning in Data Science
- Head TA CIS555: Internet and Web Systems

Skills Summary

- **Languages**: Java, Python, C++, SQL, Unix scripting
- **Tools**: Pandas, PyTorch, Numpy, CVXPy, sk-learn
- **Graduate Course Taken**: Big Data Algorithm, Modern Convex Optimization, Modern Topics in Uncertainty Estimation, Principles of Deep Learning, Computational Learning Theory, Computational Complexity, Probability Theory